

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:43:15 ; Search time 210.42 Seconds

(without alignments)

108.227 Million cell updates/sec

Title: US-09-331-631A-1

Perfect score: 3542

Sequence: 1 MAINTSNCSSLFLLSFL... SPRSTKQQOPLVSLDFVGF 666

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 3419395 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*

1: /SIDS1/gcdata/geneseq/geneseq/AA1980.DAT: *
2: /SIDS1/gcdata/geneseq/geneseq/AA1981.DAT: *
3: /SIDS1/gcdata/geneseq/geneseq/AA1982.DAT: *
4: /SIDS1/gcdata/geneseq/geneseq/AA1983.DAT: *
5: /SIDS1/gcdata/geneseq/geneseq/AA1984.DAT: *
6: /SIDS1/gcdata/geneseq/geneseq/AA1985.DAT: *
7: /SIDS1/gcdata/geneseq/geneseq/AA1986.DAT: *
8: /SIDS1/gcdata/geneseq/geneseq/AA1987.DAT: *
9: /SIDS1/gcdata/geneseq/geneseq/AA1988.DAT: *
10: /SIDS1/gcdata/geneseq/geneseq/AA1989.DAT: *
11: /SIDS1/gcdata/geneseq/geneseq/AA1990.DAT: *
12: /SIDS1/gcdata/geneseq/geneseq/AA1991.DAT: *
13: /SIDS1/gcdata/geneseq/geneseq/AA1992.DAT: *
14: /SIDS1/gcdata/geneseq/geneseq/AA1993.DAT: *
15: /SIDS1/gcdata/geneseq/geneseq/AA1994.DAT: *
16: /SIDS1/gcdata/geneseq/geneseq/AA1995.DAT: *
17: /SIDS1/gcdata/geneseq/geneseq/AA1996.DAT: *
18: /SIDS1/gcdata/geneseq/geneseq/AA1997.DAT: *
19: /SIDS1/gcdata/geneseq/geneseq/AA1998.DAT: *
20: /SIDS1/gcdata/geneseq/geneseq/AA1999.DAT: *
21: /SIDS1/gcdata/geneseq/geneseq/Aa2000.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	3542	100.0	666	19 W62828 Macadamia integrifolia
2	3412	96.3	666	19 W62829 Macadamia integrifolia
3	3227	91.1	625	19 W62830 Macadamia integrifolia
4	1127	31.8	566	13 R220181 Sequence encoded by Gossypium hirsutum
5	1103.5	31.2	590	19 W62832 Theobroma cacao an
6	1025	28.9	525	19 W62831 Zea mays antinmicro
7	898	25.4	593	19 W62835 Hordeum vulgare an
8	864.5	24.4	637	19 W62837 Peanut allergen, A
9	849.5	24.0	626	20 V15244 Peanut allergen, A
10	849	24.0	626	20 V25657 Peanut allergen, A
11	849	24.0	614	18 W22149 Peanut allergen, A
12	849	24.0	614	19 W62834 Arachis hypogaea a

ALIGMENTS				
RESULT	1	W62828	ID	W62828 standard; Protein; 666 AA.
XX	XX	AC	W62828;	XX
XX	XX	DT	27-OCT-1998 (first entry)	XX
DE	Macadamia integrifolia antimicrobial protein; infestation; control.	KW		XX
OS	Macadamia integrifolia.	XX		XX
XX	Key	FH	Location/Qualifiers	
FT	Peptide	FT	/note= "signal peptide"	
FT	Protein	FT	/note= "mature protein"	
XX	PN	W09827805-A1.		
XX	PD	02-JUL-1998.		
XX	PF	22-DEC-1997;	97WO-AU00874.	
XX	PR	20-DEC-1996;	96AU-0004275.	
XX	(RETR.) COOP RES CENT TROPICAL PLANT PATHOLOGY.			
XX	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;			
XX	WPI: 1998-377279/32.			
DR	N-PSDD; VA2310.			

Paulo

PT	Novel anti-microbial protein from e.g. Macadamia integrifolia -	DE	Macadamia integrifolia antimicrobial protein.
PT	useful for controlling microbial infestations of plants or mammals	XX	antimicrobial protein; infestation; control.
PS	Claim 1; Page 34-36; 96pp; English.	XX	
XX	The sequence is that of an antimicrobial protein which can	OS	
CC	be used to control microbial infestations in plants and mammalian	Macadamia integrifolia.	
CC	animals.		
XX	Sequence 666 AA;		
SO			
Query Match	100.0%; Score 3542; DB 19; Length 666;		
Best Local Similarity	100.0%; Pred. No. 2e-299;		
Matches	666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 MAINTSNLCSLFLSLFLSLFLSTVSLAESFEDRQEYECKRQCMOLETSQMRRCVSQCD 60	XX	
Db	1 maintsnlcslflslflslflsttvsiaeasdfrqeyeeckrqcmoletsqmrrcvsqcd 60	XX	
OY	61 KRFEEDIDWSKYDNDQEDPOTECQOCQQRRCROEQESPRQQYCQRRCKECEEEERYNQR 120	XX	
Db	61 krfeedidwskydndpotecqocqrccroeqesprqqycqrckelceeeeynqr 120	XX	
OY	121 DPQQQEYEQCQHQCREEPEERPHMOCQRCRERRYKEKRKQKREEQDEEYERM 180	XX	
Db	121 dpqqyeqeqckhqcqkrcqrecephtmqtcqqrccryekrkqkryeqeqdckyerm 180	XX	
OY	181 KEEDDKRDPDQPREYCRARRCEQQPRQHQOCOLCREQDQRCRQHQGCGDMNPORGSGRY 240	XX	
Db	181 keeddkrdpdpqreyadcrarrceeqeqprqgqrcqrcqesprqqycqrckelceeeeynqr 240	XX	
OY	241 EEEGEQSDNPYFFERSLSLSTRFEEGHISYLEMFYGRSKLRLALKNYKVLLEANPNA 300	XX	
Db	241 eegeeqsdnpypferslsistrfteeqhsivlenfygrskllrlalknyrlvileanpna 300	XX	
OY	301 FVLPLHLDADAILVGGRAKMLHDNRESYNIECGDVIRIPAGTFLINRNNERL 360	XX	
Db	301 fvlpthldadailvlggalkmlhdnresyniecgdviripagtflinrnnrl 360	XX	
OY	361 HIAKFQQTSPGQYKEFPAGGONPEPYLSTSEKILEALNTOTEKLRGVEGQREGV 420	XX	
Db	361 hiakfqtspgqykeffagggonepepylstfskellealaantqaerirvgqregv 420	XX	
OY	421 ITRASQEQLRLTRDSESHWHIRRGESSRGPVNLFNRPLLYSNKIGAYEVKPEDYR 480	XX	
Db	421 irasqeqlrltrdsehswhirrgessrgrpvnlfkrlynskyqayevkpedyr 480	XX	
OY	481 QLQDMGLSVITANVQGSMMPFFNFRSTKVVVAWSGEADVEMAPHLUSORHGGGGKR 540	XX	
Db	481 qlqdmglsvianvqgsmmpffnfrstkvvvavasgeadvemaphlusrhggggkr 540	XX	
OY	541 HEEEEVHYCOVARLUSKREAIWVAGHPVFVSSGENILFAGINGAQNHHENFLAGR 600	XX	
Db	541 heeedvhhyeqvarliskreaiwvaghpvvfssgenilfafqinagnhhenglqr 600	XX	
OY	601 ERNVHQIQEQQAMELAFAARPKKEYEESFNQDQSIFFPSPRQHQOQSPTKQQPLVSI 660	XX	
Db	601 ernvhqiqeqamelefakrkeyeesfnqdsiffpqrphqgqsprtqggplvsi 660	XX	
OY	661 LDFVGF 666	XX	
Db	661 ldfvgf 666	XX	
RESULT	2		
ID	W62829 standard; Protein; 666 AA.		
XX			
AC	W62829;		
XX			
DT	27-OCT-1998 (first entry)		
XX			

Db	421 iiasqeqirelrrddsserrwhirrgesssqgynlfnkryqgavekpedyr	480	QY	102 COPRCKEICESEEEFYNRQDPOOQYEQCQKHCORRETEPRINOTCORCERYEKRKQ	161
OY	481 QLOQMDLSVFIANTVQGSMGPPFNTRSTKVWVASEDAVMACPHLSGRIGRGSGKR	540	Db	61 cqcrckeiceeeeynqrdrpqeqyecqkccrqreteprinmcqgqrceryekrkq	120
Db	481 qldmdvsfianitqsgsmgppfrstkvvvvasgeadveimacphlsgrigrggkr	540	QY	162 QRYEEQOREDEEKKYERMKEDDKRQDQQREYEDERRRCQEQPSQHQOCQLRCRQR	221
QY	541 HEEEDWHDHYEQYRARLISKREATIVLWLAGHPVVFTSGMENLLFAFGINAQNHNENFLAGR	600	Db	121 qkryeedqredeekyeermkengdkrpkqreyedorrhceqeqpqrqccqeqqr	180
Db	541 heeedvhyyeqykariskreativlpvpyqppvfvsgnenllfafqinagnenflagr	600	QY	222 QHGGDMMNPORGSSGYEEQSDNPYFDERSLSTTRTECHISILENFVGRSK	281
QY	601 ERNULOQTEPQAMELAAPRKEVEESFSNQDQSIFPPGPROHQOSPRSTKQQPLVSI	660	Db	181 q1grrggalmporggsqgyeqegeeksdnpfyderslstrteeghislenfgsrk	240
QY	661 LDPFGF 666		QY	282 LRALKNYRLVLEANPNAYPHTIDADATILVIGRGAJKMHNDNRESYNLEAGDV	341
Db	661 ldpvgf 666		Db	241 lralknyrlvleanpnayphtidadatilviggakmihdnresynlecgvi	300
RESULT	3		QY	342 RIFAGTFYLTDNNRHLAKFLDTISTGQYKRFPGQNPRLYPLTSKELA	401
W62830	W62830 standard; Protein; 625 AA.		Db	301 r1pagttrfy1ndnnrhlakflqtlstpqykeffpagqgnpepylistfskeileaa	360
ID	W62830		QY	402 INTQTEKLURGVQREGVIRASQEOTRELTDDBSERHHRHRRGESSGPYLNKR	461
XX			Db	421 p1ysnkyqgavekpedyrq1qdmvsfianitqsgsmgppfrstkvvvvasgead	480
AC	W62830;		QY	522 ENACPHLSGRIGRGSCKRHEEEDWHDHYEQYRARLISKREATIVLWLAGHPVVFTSGMENLL	581
XX			Db	481 emacphlsgring999krheeeeevyeyqykariskreativlpvpyqppvfvsgnenll	540
DT	27-OCT-1998 (first entry)		QY	582 LEAFGINAQHNHENFLAGRERWVQLEPQAMELAAPRKEVEESFSNQDQSIFPPGPR	641
XX			Db	541 lfafqinagnenflagrenq1qeqpamelataasrievveelfnsqdesifppr	600
DE	Macadamia integrifolia antimicrobial protein.		QY	642 QHQQOSPRSTKQQPLVSIIDPVGF 666	
XX			Db	601 qhqqqsprstkkqqplvsi1dfvgf 625	
KW	antimicrobial protein; infestation; control.		RESULT	4	
OS	Macadamia integrifolia.		R20181	102 COPRCKEICESEEEFYNRQDPOOQYEQCQKHCORRETEPRINOTCORCERYEKRKQ	161
XX			ID	R20181 standard; Protein; 566 AA.	
FH	Key Peptide		XX	R20181;	
FT	Location/qualifiers		DT	16-APR-1992 (first entry)	
FT	1..28 /note= "signal peptide"		DE	Sequence encoded by 67 kd T. cacao protein cDNA.	
FT	29..666 /note= "mature protein"		XX	Cocoa; flavour; vicilin; seed storage protein.	
FT			OS	Theobroma cacao.	
PN	W09827805-A1.		XX	W09119801-A.	
XX			PD	26-DEC-1991.	
PD	02-JUL-1998.		XX	PF 07-JUN-1991: 91WO-GB00914.	
PF	22-DEC-1997; 97WO-AU00874.		XX	11-JUN-1990: 90GB-0013016.	
XX			XX	(MRSC) MARS UK LTD.	
PR	20-DEC-1996; 96AU-0004275.		PA	Spencer ME, Hodge R, Deakin EA, Ashton S;	
XX			PI	WPI; 1992-024418/03.	
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.		DR	N-PSDB; Q20377.	
PA	PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;		XX	Recombinant cacao proteins - are responsible for flavour in cocoa beans and produced in large quantities using yeast and bacterial	
XX	XX WPI; 1998-377279/32.		CC		
DR	DR N-PSDB; V4316.		CC		
XX			CC		
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals		CC		
PT	Claim 1; Page 43-45; 96pp; English.		CC		
XX	The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.		CC		
PS			CC		
XX	Sequence 625 AA;		SQ		
Query Match	91.1%; Score 3227; DB 19; Length 625;				
Best local similarity	96.6%; Prep. No. 4 5e-272;				
Matches	604; Conservative 9; Mismatches 12; Indels 0; Gaps 0;				
OY	42 QCMQLETSQMRVCVSQCDKREEDIDWSKYDNQEPOTEQOCQCRRCRQESGRQQQY 101				
Db		1			
	qomqlqtsqmrvcvsqcdkrtedidwskyanqdpgqecqgcorrccqesdprqqy 60	0			

PT expression vectors
 XX
 PS Claim 4; Fig 2; 59pp; English.
 XX
 CC The inventors claim a 67 kD and 31 kD T. cacao protein, and
 fragments, and encoding DNAs. The 47 kD and 31 kD proteins are
 derived from the 67 kD precursor. T. cacao protein cDNA was
 detected in a cDNA library prepared from immature cocoa beans RNA
 using a probe based on the AA sequence of a CNBr peptide common to
 the 47 kD and 31 kD polypeptides. Homology searches revealed close
 homologies between the 67 kD polypeptide and the vicilins, which are
 seed storage proteins.
 CC
 XX SQ Sequence 566 AA;
 Query Match 31.8%; Score 1127; DB 13; Length 566;
 Best Local Similarity 40.4%; Pred. No. 1.4e-89;
 Matches 233; Conservative 109; Mismatches 177; Indels 58; Gaps 13;
 Ps
 QY 109 ICEEEBEYNR --QDRDPQQQEKCQKHCQRRETERHMQTCQQRERRYEKEKKRQQKRY 165
 Db 22 lcsqsgsaygrkqyedprqgyeqcrrcesateerreqeqcqrer-----ey 70
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 XX DR Bower NI, Gouitter KC, Green JL, Manners JM, Marcus JP;
 XX DR WPI; 1998-377279/32.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling microbial infestations of plants or mammals
 XX
 Ps Claim 1; Page 49-51; 96pp; English.
 XX CC The sequence is that of an antimicrobial protein which can
 be used to control microbial infestations in plants and mammalian
 CC animals.
 XX SQ Sequence 590 AA;
 Query Match 31.2%; Score 1103.5; DB 19; Length 590;
 Best Local Similarity 38.7%; Pred. No. 1.7e-87;
 Matches 235; Conservative 110; Mismatches 175; Indels 87; Gaps 14;
 Ps
 QY 283 LRALKWYRLVLEAPNNAFLPFLTDADATLIVGGRGALKMWHIDNRESYNLEGCDVIR 342
 Db 176 1kgqindianfanpeaptphilphobdaalyftvqgkgtftvheneeskynpvqgvtv 235
 PA 343 IPAGGTYFLINRDNNERLHAKFLQIYSTPGQKRPAGCONPPYLSTFSKELDEAL 402
 Db 236 vpagstvyyvsgndqekltiavlaipnsvrgkyleffpgnnpkpsyygafsvyletvf 295
 QY 403 NTOTETLRLGRYFGQQR -----EGTITRASQEQLRELTDRDSESRHWHRGGSSRG 454
 Db 296 ntqrekleeileeql9qkrqggqqmrrkakapeqkiraqsgatatsph--r9ge-rta 349
 QY 455 YNLFRKRPLPSNKYQAYEVKPEDRQLQMDLSFIANTYQGSAMGPFFNTRSKVWW 514
 Db 350 inllsqsgpysngndrfacpeqdfsqfqmdvavsaasklingqalifvphynskattfv 409
 QY 515 ASGEADVEMACPHLSGRHGGGGK-RHEEEDV-----HYEQVRAKLSKREAVVLA 566
 Db 410 tdygyyaqmacphlsrqssqsgqrdireeqeeseesetlgefqqyakplspgadvpa 469
 QY 567 GHPVVFVSSGENLNLLFAFGINAONHNENTLAGRNVNLOIEPQAMELAFAAPKEVEE 626
 Db 470 ghavtffaskdqphlnavafqinaqnriFlagk-knlyrqmdseakeleisrgvpsklvn 528
 QY 627 SFNSPDQSISIFGPQHQOOQSPRSIKQQQLVLSDLF 663
 Db 529 ifnnpdseyfmfsqqr---rderrgoplasildf 562
 RESULT 5
 W62832
 ID W62832 standard; Protein; 590 AA.
 AC W62832;
 XX
 DT 27-oct-1998 (first entry)
 XX DE Gossypium hirsutum antimicrobial protein.
 XX KW antimicrobial protein; infestation; control.
 XX
 OS Gossypium hirsutum.
 XX PN WO9827805-A1.
 XX PD 02-JUL-1998.
 XX PF 22-DEC-1997; 97WO-AU00874.
 XX PR 20-DEC-1996; 96AU-0004275.
 XX PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PT Bower NI, Gouitter KC, Green JL, Manners JM, Marcus JP;
 XX DR WPI; 1998-377279/32.
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling microbial infestations of plants or mammals
 XX
 Ps Claim 1; Page 49-51; 96pp; English.
 XX CC The sequence is that of an antimicrobial protein which can
 be used to control microbial infestations in plants and mammalian
 CC animals.
 XX SQ Sequence 590 AA;
 Query Match 31.2%; Score 1103.5; DB 19; Length 590;
 Best Local Similarity 38.7%; Pred. No. 1.7e-87;
 Matches 235; Conservative 110; Mismatches 175; Indels 87; Gaps 14;
 Ps
 QY 76 EDPQTECQCQCORRCROQESGRQQQCQRCRKEICEEEBEYNRDF-DPOQYEQCQKHC 133
 Db 35 ddpkkyedcrccrdtrgqkqeqgqceesccsqqgekdqqrhpdpqrreecqec 94
 QY 134 ORRETPEPRHMOTCOQCERREKEERKQKRYEEQOREDEBEKYEEPRMKEEDNKRDPQRE 193
 Db 95 -rqqeerqqpcqccrkifqeqeqq-----sqrq 123
 QY 194 YEDCRRCRCEQD-----FQOHOCOLCRCREOORQHGRRGGDMMPQRG-----GSGRYEEGE 245
 Db 124 fgecggnchqgeqrbkqgcvcrecrekyqe-----npwrgereeeaeeteaqeq 175
 QY 246 EQSDNPYVYFDRSLSSTRFREPEGHISVLEMFYGRSKLRLAKNVRVLLEANPNAFLPT 305
 Db 176 eqshphpfhfsqrsqfrfreehgnfrvlfqfashtpilqinefnrlsleanpntfvlh 235
 QY 306 RLDADAILLVGGRGALKMWHIDNRESYNLCEGVDYIRPATGTFYLINRDNERLHAK 365
 Db 236 hodaekiyitngtgrtlfttheneeskynivpgvvkvpagstvylangdnkekliavl 295
 QY 366 LQTISPGQYKEFAGGQREPEPYLSTFSEKTELEALNTQTEKURGVFG-----QORE 418
 Db 296 hrpvnnpgqfeeftpagsqpgsyirafsrleilepafntrseqldelfgrqsrrrqgg 355
 QY 419 GVIIASQEQTRELTRDSDSSRHWHRGGESSRCPYINFNRPLYSNKYQAYEVKPED 478
 Db 356 qmfkpasqedairalsqeqatpr---eksgc-rafnlslqtprysnangrfteacpe 409
 QY 479 YRQLOQMDLSFIANTYQOSMMGPFFNTASTKVVVASGBDEVEMACPHLSGRHGGGG 538
 Db 410 frqlqldinvlvtsalqlnqsgsifvphynskattfvltvegngyaemvslph----rqss 464
 QY 539 KRHEEEDV-----HYEQVRAKLSKREAVVLAFLGHPVVFSSGENLNLLFA 585
 Db 465 yeeeeeedeeeqeqeerrrsgqkirkirsrlsrgdfvwpnfptfvasqsnqlrmqf 524
 QY 586 G----INAQNHNENFLAGRERNVIQIQIPQAMELAFAAPKEVEESFNQDSIIFPGP 640
 Db 525 glynqinpdhqnrifvagknlh-rqwdqsaqkelaqyvssrlveifnsnpqesyfvs- 582
 QY 641 RQHQQS 647

XX antimicrobial protein; infestation; control.
 KW XX
 OS XX Arachis hypogaea.
 PN XX W0982705-A1.
 PD XX 02-JUL-1998.
 PR XX 22-DEC-1997; 97WO-AU00874.
 PT XX 20-DEC-1996; 96AU-0004275.
 PA XX (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 XX
 PT XX Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
 PS XX DR WPI; 1998-377279/32.
 CC XX Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling microbial infestations of plants or mammals
 XX
 PS XX Claim 1; Page 55-57; 96pp; English.
 CC XX The sequence is that of an antimicrobial protein which can
 be used to control microbial infestations in plants and mammalian
 animals.
 XX SQ Sequence 614 AA;

Query Match 24.0%; Score 849; DB 19; Length 614;
 Best Local Similarity 34.7%; Pred. No. 2; Je-65;
 Matches 219; Conservative 110; Mismatches 216; Indels 86; Gaps 21;
 QY 70 SKYDNDQEDPQTEC-OCQQRCCRQEQSCPRQQYCQQRCKECEEEFYRNORDPQOQEYQ 128
 DB 27 SPYKTCNP---CAQRLQSC-QQEPDLKQKACERCTKL---EY---DPREVYD- 72
 QY 129 CQKHCORRETEPRHMORQCRQERRYEKEKRQKAYEEOREDDEYERMEEDENKRD 188
 DB 73 ---t-gatnqrhppegtgrqpdyyddarqrprreergwgpaeprereed-wrq 125
 QY 189 PQOREVEDCRRCEQEPHQHQLRCREQORQHGRGDMMPQRGSGRYEEBEEQS 248
 DB 126 P-----EDWRRPSSHQ-PR-----KIRPEGRGE---QEWGTPGSEVREETSR 165
 QY 249 DNPYYFEDERSLSTRFRTEEGHSVLVNFYGHSLRALKNVRLVIFANPNAFVPLTHD 308
 DB 166 NNPFFYPSRSTRYQNGRFLQVLFQDFQSKQFQPLQHRIYQEARPLTVLPKHD 225
 QY 309 ADAILYVIGRGALKMTHDNRESYNEECGGVIRATPGTYFLINNDNNERLHAKFLQT 368
 DB 226 ADNLILRQGGAATVANGNPKSFLDGHAIPLPSGFISYILNHNDQPLRKVKISM 285
 QY 369 ISTPGOYKEFPAGONPEPYLSTSKELPAALNQTEKLRGVF-----GQ 416
 DB 286 VNTPGQGFEDFIPASSRDRQSSYIQQFSRNTLEAFAHAEFNIRRVLILEENAGGEGERQR 345
 QY 417 R-----EVIRASQEQIRELLRDDESRRHWHIRRGESS- RGPNLFNKAPLYSN 466
 DB 346 RRTSTSNDNEGYIVKSYWQELTHAKSVS---KKGSEEDITNPINLRDGPDSL 401
 QY 467 KYGQAEVKPEDYR-QLQMDPLSVFIANVTOGSMMGPFFNTRSTKVVVVASGEADEMAC 525
 DB 402 nfgfrfekvpkknqlqdmlmmtcveikgalnlpfnksamrnvwtktgtnleiva 461
 QY 526 PHLSGRHGGRGKRRHEEEEDVHE---OVR---ARLSKREATIVLAGHPVVFSSGN 578
 DB 462 vrkqqggrgreqeweeeedeeegsnnrsvrryarlkgdgvfinpaahpvainassel 521
 QY 579 NULLFCGINAQNNHNFLAGRERRNLQOEPQAMELAFAAPRKVEEESNSQDOSIFFP 638

RESULT 13
 ID W22150
 PE W22150 standard; Protein; 626 AA.
 AC W22150;
 DT XX 29-DEC-1997 (first entry)
 DE XX Peanut allergen Ara h1.
 KW XX Peanut; seed storage protein; allergen; allergy; hypersensitivity;
 monoclonal antibody; ELISA; analysis; Ara h1.
 OS XX Arachis hypogaea strain Florunner.
 FH Key Location/Qualifiers
 FT Peptide 1..22
 FT Protein /label= Sig_peptide
 FT Modified-site 521..523
 FT /note= "N' glycosylation site"
 PN XX W09724139-A1.
 RD XX 10-JUL-1997.
 PR XX 23-SEP-1996; 96WO-US1515222.
 PR 04-MAR-1996; 96US-0610424.
 PR 29-DEC-1995; 95US-0009455.
 PA XX (UYAR-) UNIV ARKANSAS.
 PT XX Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;
 DR XX WPI; 1997-363453/33.
 NN XX N-PSDB; T76613.
 PT XX Peanut allergens Ara h1 and Ara h1I - used for vaccination and in
 two-site monoclonal antibody based ELISA
 XX PS Claim 31; Page 172; 354pp; English.
 CC This polypeptide comprises major peanut allergen Ara h1 (W22149).
 Its sequence was deduced from cDNA clone p41b (T76613), isolated
 from peanut seed cDNA using a primer (see T76616) based on an
 isolated Ara h1 peptide (see W24206). The sequence shows
 significant homology with the vicilin family of seed storage
 proteins of other legumes. The allergen is recognised by serum
 IgE from a large proportion of individuals with peanut
 hypersensitivity. Ara h1 and Ara h1I (see W2164) can be used to
 raise monoclonal antibodies which are used in a specific two-site
 assay for the detection of Ara h1 or Ara h1I (claimed). IgE
 binding Ara h1 antigen epitopes (see W24165-87) may be used in
 vaccines to protect against allergic reactions to peanut allergens,
 e.g. anaphylactic shock.

SO Sequence 626 AA;

Query Match 23.8%; Score 843.5; DB 18; Length 626;
 Best Local Similarity 34.2%; Pred. No. 7.2e-65;
 Matches 211; Conservative 108; Mismatches 205; Indels 93; Gaps 20;

OY	93 ESGPRQQY--CORACKCCEEEEVNRQDPOQDQEQQCKHCCORETEER-----H	142	PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
Db	26 ksspykktcnpcacqlcqscqepdalkqk-----acesrctkleydpbcvprgh	78	useful for controlling microbial infestations of plants or mammals
OY	143 MOTCQRC--ERR-----YEKERKQQRVEOREDEBEKYEERMEKEEDNKRDPQR	192	XX PS Claim 1; Page 63-65; 96pp; English.
Db	79 tqttngrrsppegtgrgpgyddfr--qprredeggrwpgapgrerered-wqrpr--	133	CC The sequence is that of an antimicrobial protein which can
OY	193 EYEDCRRCQEOPRQHQCLRCREQORHQGRGGMNPORGGSARVEEEDSDNPY	252	CC be used to control microbial infestations in plants and mammalian
Db	134 --edwrlrpshq-p-----kirpegreg---qewgtpgshvreets-nupf	175	XX animals.
OY	253 YFDERSLSTRETECHISLVIENFYRSKLLRALKNVRVLEAENNAFVPLTHIDATI	312	Sequence 605 AA:
Db	176 yfpstrfstrsynqnqngirvqrgfdarsrqsnqnlqnlrvleakptlvipkhadni	235	Query Match 23.7%; Score 840.5; DB 19; Length 605;
OY	313 LLVIGRGALAKMHNHNRSTNLNECDVIRPAGFTYLINRDNERLHIAKFLQTISTP	372	Best Local Similarity 30.5%; Pred. No. 1.3e 64; Matches 205; Conservative 144; MisMatches 223; Indels 101; Gaps 17;
Db	236 lviqggatvtvngnkrksfldeghaitspsgfsiylnrhdqnlrvakiqmvptq	295	QY 11 LIFLISFLSLSTVS-LAESEFDROEYEECKROCMOLETSQMRRCVSODDKRFEDID 68
OY	373 GOKEFFPAGGONPEPLSTFSKELIEALNTQTEKLRGVF-----GQOR--	417	Db 8 llilglivflasvsfsgiyawekenkphnkclqschserdsyrmqacharc-----n 59
Db	296 ggffeflpasardqssylqgsrsntleaaafnaefneirrvilleenaggegeergarwst	355	QY 69 WSKYDNQEDPOTECQCQCRQRQEQSGPQQYCQRRCKETCEEEVYNRORD--PQQ 125
OY	418 ----EGVIRASQOIRETRDSDSESRWHIRGESSRG----PYNLFNKPLYSNK 467	470	Db 60 llkvkeecegeipprprpqhpeprpqge-----keededapripipfprq 110
Db	356 rsseennegvivkvskehveeltkaks---vskbgseeedgdtipinlregedpslnn	410	QY 126 YEQCKQKICORETE-PRHMQTCQRCERRYKEKKRKQKRYEEQDREDEBEKYEERMEKEED 184
OY	468 YGQAVAYEVKPEDYR-QQDMDISVFYANVTOGSMGGFFNRSTKVWVVAEGADEMACP	526	Db 111 prqeeheehqreqewpr-----keekrgekgsseededebeqder--qfp 154
Db	471 rkeqqgrg----reeeededeeegsnevrtytaikqdfvimpahpvainssell	527	QY 185 NKRDPOREYIDCCRKEQQPQQHOCQLCREQORHQGRGGMNPORGSGRVEE 244
OY	580 LLLFARGINAQNHNENFLAGRERNVQIPEQAMLAFAAPRKEVEESNSQDSIFFPG	639	Db 155 fprpphqke-----erneeedede-----eqgresesed-----selr 188
Db	528 11-gfginaenhriflagkdvnidqiekqakdlafpsgsqeqeklikngkeshfvs	585	QY 245 ERQSDNPYFFERSLSTRFRTEEGHTSVLENFYGRKLRLAKNVRVLNLPNAFVLP 304
OY	640 PROHOOSPRSTKQOP 656	1	Db 189 rkknkoplfqsnrfelblkqygrirlqfingqspqqlqnlrdyfrlefnskpntllp 248
Db	586 rrpqsqsdspsspkesp 602	1	QY 305 THDADALILUIGGRGALKMHNHNEBSYNEECGDIVTRAGFTYLINRDNERLHIAK 364
RESULT	14	1	Db 249 nhadadvilivingtalslsvnnnddsyrlqsgdalrvpsgttlyvvnpndnenrlit 308
W62838	W62838 standard; Protein; 605 AA.	1	QY 365 FLOTISPGQKERRQEFFAGQGQPEPVUSTESKEILEALNTQTEKLRGVF-----GQ 415
AC	W62838;	1	Db 309 laipvnkpgrefestissteegqsyiaggfsmileasydkfeinkvlfrreeggqge 368
DT	27-OCT-1998 (first entry)	1	QY 416 QR-EGVIRASQEQIRETRDSDSESRWHIRGESSRGPSYNYLNKPLYSNKVGQAYE 473
XX	Glycine max antimicrobial protein.	1	Db 369 qrlqesiveskqekralskrass---rktsisedkfplrsdrpdykqskigkffe 424
XX	antimicrobial protein; infestation; control.	1	QY 474 VPKPEDYRQLOQMDLSFTIANYNGSMGPFNTSRKVVVAVSGEADVNACPHLSGRIG 533
XX	22-DEC-1997; 97WO-AU00874.	1	Db 425 itpeknqlrdldifisivdmegallphnfskaivilvimeganieiwl----- 475
PR	20-DEC-1996; 96AU-0004275.	1	QY 534 GRGGGK----RHEEEDDVHREQRARLSKREKAIYLAGHPPVFWSSGNINLLPAFGIN 588
PD	02-JUL-1998.	1	Db 476 ---gikeqqeqgqeqqeqpievkyraelseddfivpagypvvv--natsnlnraigin 530
XX	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	1	QY 589 AQNNHENFLAGGERNVLQQEPQAMLAFAAPRKEVEESNSQDSIFFGCPHQQQSP 648
PP	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;	1	Db 531 aennqnqlagsqdndvisqspqqlqsaqekliknqresyfdaqpkkkegn 590
PN	W09827805_A1.	1	QY 649 RSTKQOOLPLVLS 661
XX	22-DEC-1997; 97WO-AU00874.	1	Db 591 kgirk--gplassil 601
RESULT	15	1	PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PA	ID Y40999 standard; protein; 605 AA.	1	useful for controlling microbial infestations of plants or mammals
XX	AC Y40999;	1	XX PS Claim 1; Page 63-65; 96pp; English.
XX	06-DEC-1999 (first entry)	1	CC The sequence is that of an antimicrobial protein which can
DI	XX	1	CC be used to control microbial infestations in plants and mammalian

